Aust. N. Z. J. Stat. **44**(4), 2002, 439–446

THE ASYMPTOTIC DISTRIBUTION OF THE S-GINI INDEX

RIČARDAS ZITIKIS[1]* AND JOSEPH L. GASTWIRTH[2]

University of Western Ontario and the Laboratory for Research in Statistics and Probability; George Washington University and the National Cancer Institute

Dedicated to Professor David Vere-Jones on the occasion of his 65th birthday

Summary

Several generalizations of the classical Gini index, placing smaller or greater weights on various portions of income distribution, have been proposed by a number of authors. For purposes of statistical inference, the large sample distribution theory of the estimators of those measures of economic inequality is required. The present paper was stimulated by the use of bootstrap by Xu (2000) to estimate the variance of the estimator of the S-Gini index. It shows that the theory of L-statistics (Chernoff, Gastwirth & Johns, 1967; Shorack & Wellner, 1986) makes possible the construction of a consistent estimator for the S-Gini index and proof of its asymptotic normality. The paper also presents an explicit formula for the asymptotic variance. The formula should be helpful in planning the size of samples from which the S-Gini index can be estimated with a prescribed margin of error.

Key words: asymptotic normality; consistency; economic inequality; exponential distribution; Gini index; income distribution; L-statistic; Pareto distribution; S-Gini index.

1. Introduction

A substantial literature has been devoted to the construction of indices of economic inequality that are consistent with axiomatic systems of fairness. We refer readers to Sen (1997), Blackorby, Bossert & Donaldson (1999), Cowell (1999), Giorgi (1999), Ryu & Slottje (1999) and Silber (1999) for a survey of the area and main references.

This paper was stimulated by the use of bootstrap by Xu (2000) to estimate the asymptotic variances of generalized Gini indices. We focus on the S-Gini index

$$I_{F,\nu} = 1 - \frac{\nu}{\mu}\int_0^1 F^{-1}(t)(1-t)^{\nu-1}\,dt \qquad (\nu > 0)$$

Received June 2001; accepted July 2001.

*Author to whom correspondence should be addressed.

[1] Dept Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada N6A 5B7 and Laboratory for Research in Statistics and Probability, Carleton University, and University of Ottawa, Ottawa, Ontario, Canada K1S 5B6. e-mail: zitikis@stats.uwo.ca.

[2] Dept of Statistics, George Washington University, Washington DC 20052, USA, and Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892-2580, USA.

This paper was presented at the symposium in honour of Professor David Vere-Jones's 65th birthday in April 2001 at Victoria University of Wellington in New Zealand.

Acknowledgments. The authors thank the Technical Editor for valuable suggestions that helped improve the presentation of the paper. They also thank Stephen Donald for sending them a copy of Barrett & Donald (2000). This research was partially supported by an NSERC of Canada individual research grant at the University of Western Ontario, and by a grant from the National Science Foundation at the George Washington University.

and its large sample estimation. In the definition of $I_{F,\nu}$ above, the parameter ν is a fixed positive number, F denotes the cumulative distribution function (cdf) of a random variable X, F^{-1} is the corresponding quantile function and $\mu = \mathrm{E}(X)$ denotes the mean of X which is assumed to be finite and non-zero. Later, we also assume the finiteness of some higher order moments of X, but no continuity-type assumptions are imposed on F. Thus, the results of this paper are applicable to all cdfs, including those that are continuous and discrete.

The S-Gini index $I_{F,\nu}$ is well defined for some classes of cdfs F, depending on the value of ν. Yitzhaki (1983) subdivides the positive values of ν into three regions by noting that the values $0 < \nu < 1$ reflect equality aversion, the value $\nu = 1$ reflects equality neutrality and values $\nu > 1$ reflect inequality aversion. In the case of inequality aversion ($\nu > 1$) the index $I_{F,\nu}$ is always well defined because the first moment μ is finite by assumption. In the case of equality neutrality ($\nu = 1$) the index $I_{F,\nu}$ is identically 0. For the equality aversion case ($0 < \nu < 1$) the index $I_{F,\nu}$ is well defined provided that the rth absolute moment $\mathrm{E}(|X|^r)$ is finite for some $r > 1/\nu$. These moment assumptions describe the range of parameters for which a large sample estimation theory for $I_{F,\nu}$ is possible and, indeed, needed.

Several approaches have been proposed and used for estimating the S-Gini and related indices of economic inequality. In particular, Mehran (1976), Nygård & Sandström (1989), Giorgi & Pallini (1990) use the theory of L-statistics in their studies. Barrett & Donald (2000) use the theory of empirical and quantile processes to study the S-Gini and other indices of inequality and poverty.

In this paper we demonstrate that L-statistics provide a most suitable and convenient technical tool for developing a large sample estimation theory for the S-Gini index $I_{F,\nu}$. For this reason we now express the S-Gini index $I_{F,\nu}$ in the form

$$I_{F,\nu} = 1 + \frac{1}{\mu} T_{F,\nu}, \quad \text{where} \quad T_{F,\nu} = \int_0^1 F^{-1}(t)\, d\Psi_\nu(t) \quad \text{and} \quad \Psi_\nu(t) = (1-t)^\nu .$$

When $\nu = 1$, $T_\nu(t) = -\mu$, and thus $I_{F,\nu} = 0$. When $\nu = 2$, $I_{F,\nu}$ is the classical Gini index

$$G_F = \frac{1}{2\mu} \mathrm{E}(|X_1 - X_2|),$$

where X_1 and X_2 are independent random variables with cdf F. The quantity $T_{F,\nu}$ is the asymptotic expectation or, in other words, the centring constant of the L-statistic (see e.g. Chernoff *et al.*, 1967; Shorack & Wellner, 1986),

$$T_{n,\nu} = \sum_{i=1}^{n} \left(\Psi_\nu\Big(\frac{i}{n}\Big) - \Psi_\nu\Big(\frac{i-1}{n}\Big) \right) X_{i:n},$$

where $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ are order statistics of random variables $X_1, X_2, \ldots, X_n$. Consequently, the S-Gini index $I_{F,\nu}$ is estimated by the ratio

$$I_{n,\nu} = 1 + \frac{1}{\bar{X}} T_{n,\nu},$$

where $\bar{X}$ denotes the sample mean of $X_1, X_2, \ldots, X_n$. Using the definition of Ψ_ν, we can now rewrite the estimator $I_{n,\nu}$ as follows:

$$I_{n,\nu} = 1 - \frac{1}{\bar{X} n^\nu} \sum_{i=1}^{n} \big((n-i+1)^\nu - (n-i)^\nu \big) X_{i:n}.$$

The latter form of the S-Gini estimator $I_{n,\nu}$ is the one that is most frequently used in the literature (see e.g. Weymark, 1980/1981; Chakravarty, 1988; Blackorby *et al.*, 1999; Xu, 2000). When $\nu = 1$, $T_{n,\nu} = -\bar{X}$, and thus $I_{n,\nu} = 0$. When $\nu = 2$, $I_{n,\nu}$ is the classical Gini index estimator (see e.g. David, 1968, 1970 and references therein),

$$G_n = \frac{1}{2n^2\bar{X}} \sum_{i=1}^{n}\sum_{j=1}^{n} |X_i - X_j| .$$

Using the estimator $I_{n,\nu}$, we develop a large sample estimation theory for the S-Gini index $I_{F,\nu}$ aiming at optimal assumptions on the underlying cdf F.

The paper is organized as follows. In Section 2 we consider strong consistency of the S-Gini estimator $I_{n,\nu}$. The main result is formulated in Theorem 1. It says, loosely speaking, that $I_{n,\nu}$ is a strongly consistent estimator of $I_{F,\nu}$ for any value of $\nu > 0$, whenever $I_{F,\nu}$ is finite. In Section 3 we consider the asymptotic normality of $I_{n,\nu}$. The main result is formulated in Theorem 2 where we show that the S-Gini estimator $I_{n,\nu}$ is asymptotically normal for the values $\nu > \frac{1}{2}$, provided that some moment assumptions hold. In Theorem 2 we also give an explicit formula for the asymptotic variance of the appropriately centred and normalized $I_{n,\nu}$. Section 4 illustrates the results of Sections 2 and 3 when the cdf F is exponential or Pareto. We calculate a number of quantities that appear in Sections 2 and 3 for the two distributions and present the results in Table 1. The results and their counterparts for other distributions can be used to plan the sample size needed to estimate the S-Gini index with a prescribed margin of error. The results of Section 4 in the Pareto case show that the assumptions of Theorems 1 and 2 cannot be improved upon for arbitrary (i.e. general) cdfs. In Section 5 we discuss approaches for estimating the asymptotic variance of $I_{n,\nu}$. These results are useful when constructing confidence intervals for the S-Gini index $I_{F,\nu}$.

2. Strong consistency of $I_{n,\nu}$

We prove that the S-Gini estimator $I_{n,\nu}$ is strongly consistent under the same assumptions on X that were given in Section 1 when the existence and finiteness of $I_{F,\nu}$ were discussed.

Theorem 1. *If $0 < \nu < 1$, we assume that $\mathrm{E}(|X|^r) < \infty$ for some $r > 1/\nu$. If $\nu \geq 1$, we assume that $\mathrm{E}(|X|) < \infty$ (this assumption is automatically satisfied because the mean μ is finite). Under these assumptions, the S-Gini estimator $I_{n,\nu}$ is a strongly consistent estimator of $I_{F,\nu}$.*

Proof. Under the moment assumption $\mathrm{E}(|X|) < \infty$ the sample mean $\bar{X}$ is a strongly consistent estimator of μ. Thus, the theorem follows if we show that $T_{n,\nu}$ is a strongly consistent estimator of $T_{F,\nu}$. To prove the latter, we use Shorack & Wellner (1986 Theorem 5 pp. 666–667). Having verified the assumptions of that theorem, we obtain that $T_{n,\nu}$ is a consistent estimator of $T_{F,\nu}$ provided that $\mathrm{E}(|X|^r) < \infty$ for some r such that $r > 1/\nu$. When $\nu = 1$, the latter assumption requires the finiteness of $\mathrm{E}(|X|^r)$ for some r such that $r > 1$. However, when $\nu = 1$ we need only $\mathrm{E}(|X|) < \infty$, since $T_{n,1} = -\bar{X}$, $T_{F,1} = -\mu$ and $\bar{X}$ is a strongly consistent estimator of μ under the assumption $\mathrm{E}(|X|) < \infty$.

3. Asymptotic normality of $I_{n,\nu}$

We show that $\sqrt{n}(I_{n,\nu} - I_{F,\nu})$ has an asymptotically normal distribution for any $\nu > \frac{1}{2}$, provided that some moment conditions are satisfied. Section 4 shows that the assumption

$\nu > \frac{1}{2}$ cannot be relaxed in the Pareto case, and thus cannot be relaxed for the general case (i.e. arbitrary cdf F) either.

Theorem 2. *If $\frac{1}{2} < \nu < 1$, we assume that $\mathrm{E}(|X|^r) < \infty$ for some $r > 1/(\nu - \frac{1}{2})$. If $\nu \geq 1$, we assume that $\mathrm{E}(X^2) < \infty$. Under these assumptions, $\sqrt{n}(I_{n,\nu} - I_{F,\nu})$ has an asymptotically normal distribution with mean 0 and variance*

$$\sigma^2_{F,\nu} = \frac{1}{\mu^2}\big(\sigma_F(\nu,\nu) + 2(I_{F,\nu} - 1)\sigma_F(1,\nu) + (I_{F,\nu} - 1)^2\sigma_F(1,1)\big),$$

where, for the values $\alpha, \beta \in \{1, \nu\}$ and notation $s \wedge t = \min(s,t)$,

$$\sigma_F(\alpha,\beta) = \alpha\beta \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \big(F(x\wedge y) - F(x)F(y)\big)\big(1 - F(x)\big)^{\alpha-1}\big(1 - F(y)\big)^{\beta-1}\,dx\,dy\,.$$

To apply Theorem 2 for constructing confidence intervals for the S-Gini index $I_{F,\nu}$ based on $I_{n,\nu}$, we need a consistent estimator of $\sigma^2_{F,\nu}$. Such an estimator is given in Section 5 where we also derive a confidence interval for $I_{F,\nu}$.

Proof. We start the proof with the representation

$$\begin{aligned}\sqrt{n}\,(I_{n,\nu} - I_{F,\nu}) &= \frac{1}{\mu}\sqrt{n}\,(T_{n,\nu} - T_{F,\nu}) - \frac{T_{F,\nu}}{\mu^2}\sqrt{n}\,(\bar{X} - \mu) \\ &\quad - \Big(\frac{1}{\mu^2\sqrt{n}}\Big)\sqrt{n}\,(T_{n,\nu} - T_{F,\nu})\sqrt{n}\,(\bar{X} - \mu) + \Big(\frac{T_{n,\nu}}{\bar{X}\mu^2\sqrt{n}}\Big)n(\bar{X} - \mu)^2. \qquad (1)\end{aligned}$$

This reduces our investigation of the asymptotic behaviour of $\sqrt{n}(I_{n,\nu} - I_{F,\nu})$ to those concerning $\sqrt{n}(T_{n,\nu} - T_{F,\nu})$ and $\sqrt{n}(\bar{X} - \mu)$. Now $\sqrt{n}(\bar{X} - \mu) = -\sqrt{n}(T_{n,1} - T_{F,1})$, and so to derive the desired asymptotic result for $\sqrt{n}(I_{n,\nu} - I_{F,\nu})$ we need only know the asymptotic behaviour of $\sqrt{n}(T_{n,\nu} - T_{F,\nu})$. For this reason we use Shorack & Wellner (1986 Theorem 5(i) p. 666). A routine verification of the assumptions of that theorem shows that, for any $\nu > \frac{1}{2}$, the representation

$$\sqrt{n}(T_{n,\nu} - T_{F,\nu}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_{i,\nu} + R_n \qquad (2)$$

holds, where

$$\begin{aligned}Y_{i,\nu} &= -\int_{-\infty}^{\infty} \big(I\{X_i \leq x\} - F(x)\big)\Psi'_\nu\big(F(x)\big)\,dx \\ &= \nu \int_{-\infty}^{\infty} \big(I\{X_i \leq x\} - F(x)\big)\big(1 - F(x)\big)^{\nu-1}\,dx\,,\end{aligned}$$

and the remainder term $R_n \xrightarrow{\mathrm{p}} 0$ in probability when $n \to \infty$ provided that $\mathrm{E}(|X|^r) < \infty$ for some $r > 2/(2\nu - 1)$. The case $\nu = 1$ needs special attention. The result above states that if $\nu = 1$, then (2) holds true with $R_n \xrightarrow{\mathrm{p}} 0$ if $\mathrm{E}(|X|^r) < \infty$ for some $r > 2$. The latter moment assumption appears to be superfluous because in the case $\nu = 1$ the remainder term R_n is identically 0. This is seen from the equality

$$T_{n,1} - T_{F,1} = \frac{1}{n}\sum_{i=1}^{n} Y_{i,1}\,, \qquad (3)$$

which follows from $T_{n,1} - T_{F,1} = -\bar{X} + \mu$ and also from the representation (see e.g. Shorack, 2000 p. 116)

$$-X_i + \mu = \int_{-\infty}^{\infty} \big(I\{X_i \le x\} - F(x)\big)\, dx .$$

When proving the asymptotic normality of $\sqrt{n}\,\big(T_{n,1} - T_{F,1}\big)$ we assume $\mathrm{E}(X^2) < \infty$, but this assumption is less stringent than the requirement $\mathrm{E}(|X|^r) < \infty$ for some $r > 2$. Applying both representations (2) and (3) to the right-hand side of (1), we derive

$$\sqrt{n}(I_{n,\nu} - I_{F,\nu}) = \frac{1}{\mu}\bigg(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_{i,\nu}\bigg) + \frac{I_{F,\nu}-1}{\mu}\bigg(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_{i,1}\bigg) + R_n' , \qquad (4)$$

where the remainder term $R_n' \xrightarrow{\mathrm{P}} 0$ when $n \to \infty$ under the assumptions of Theorem 2. From (4) we conclude that $\sqrt{n}\,(I_{n,\nu} - I_{F,\nu})$ converges to a normal random variable with mean 0 and variance

$$\sigma^2_{F,\nu} = \frac{1}{\mu^2}\mathrm{E}\Big(\big(Y_{1,\nu} + (I_{F,\nu} - 1)Y_{1,1}\big)^2\Big) .$$

Under the assumptions of Theorem 2, and using Shorack & Wellner (1986 Lemma 1 p. 663), we verify that the variance $\sigma^2_{F,\nu}$ is finite and the formula stated in Theorem 2 holds. We also note that $\sigma^2_{F,\nu} = 0$ when $\nu = 1$. This is not surprising because both $I_{n,1}$ and $I_{F,1}$ are equal to 0.

4. The exponential and Pareto cases

For several parametric families, the asymptotic mean $I_{F,\nu}$ of the S-Gini estimator $I_{n,\nu}$ and the asymptotic variance $\sigma^2_{F,\nu}$ of $\sqrt{n}\,(I_{n,\nu} - I_{F,\nu})$ can be computed and expressed in terms of parameters of the corresponding cdf F. We illustrate this for the exponential distribution with cdf

$$F(x) = 1 - e^{-\lambda(x - x_0)} \qquad (x \ge x_0 \ge 0;\ \lambda > 0) ;$$

and the Pareto distribution with the cdf

$$F(x) = 1 - \Big(\frac{x_0}{x}\Big)^{\lambda} \qquad (x \ge x_0 > 0;\ \lambda > 0) .$$

For these two distributions, Table 1 presents explicit formulae for quantities that appear in Sections 2 and 3 and, in particular, in the definition of $\sigma^2_{F,\nu}$. Using these, we can obtain, for example, formulae for the asymptotic variance $\sigma^2_{F,\nu}$ in terms of the parameters of the two distributions.

In Table 1 the regions of possible values of λ and ν are given beside each formula only if they are smaller than $\lambda > 0$, the region where the exponential and Pareto distributions are defined. In the exponential case all the quantities of Table 1 exist and are finite for any value of $\lambda > 0$ because exponential variables have finite moments of all orders. The situation is noticeably different in the Pareto case because the number of finite moments depends on the value of $\lambda > 0$. We now discuss this case in detail.

We see from Table 1 that the formula for the S-Gini index $I_{F,\nu}$ holds true in the Pareto case only when

$$\lambda > \max\Big(1,\ \frac{1}{\nu}\Big) . \qquad (5)$$

TABLE 1

Formulae for the parameters of the asymptotic distribution of $I_{n,\nu}$

Parameters	Exponential	Pareto
μ	$\dfrac{1+\lambda x_0}{\lambda}$	$\dfrac{x_0\lambda}{\lambda-1}$ $[\lambda > 1]$
$I_{F,\nu}$	$\dfrac{\nu-1}{\nu(1+\lambda x_0)}$	$\dfrac{\nu-1}{\lambda\nu-1}$ $[\lambda > \max(1, 1/\nu)]$
$\sigma_F(1,1)$	$\dfrac{1}{\lambda^2}$	$\dfrac{x_0^2\lambda}{(\lambda-1)^2(\lambda-2)}$ $[\lambda > 2]$
$\sigma_F(1,\nu)$	$\dfrac{1}{\nu\lambda^2}$	$\dfrac{x_0^2\nu\lambda}{(\lambda-1)(\nu\lambda-1)(\nu\lambda-2)}$ $[\lambda > \max(1, 2/\nu)]$
$\sigma_F(\nu,\nu)$	$\dfrac{1}{(2\nu-1)\lambda^2}$ $[\nu > \frac{1}{2}]$	$\dfrac{x_0^2\nu^2\lambda}{(\nu\lambda-1)^2((2\nu-1)\lambda-2)}$ $[\lambda > 1/(\nu-\frac{1}{2}),\ \nu > \frac{1}{2}]$

The regions of the values of λ and ν where the corresponding formulae are valid are given beside each formula, provided the regions are smaller than $\lambda > 0$ and $\nu > 0$.

When $0 < \nu < 1$, (5) is equivalent to $\lambda > 1/\nu$. If we reformulate this assumption in terms of the moments of X, we see that it is equivalent to the finiteness of $\mathrm{E}(|X|^r)$ for some $r > 1/\nu$. When $\nu \geq 1$, (5) is equivalent to $\lambda > 1$. In terms of the moments of X, this is equivalent to $\mathrm{E}(|X|) < \infty$. It follows that the assumptions of Theorem 2 are optimal in the Pareto case and thus cannot be improved upon in general.

Examination of the Pareto case shows that the assumptions of Theorem 2 cannot be relaxed. From Table 1 it follows that to calculate $\sigma^2_{F,\nu}$ we have to assume that

$$\lambda > \max\left(2,\ \frac{2}{\nu},\ \frac{1}{\nu-\frac{1}{2}}\right) \qquad \text{and} \qquad \nu > \tfrac{1}{2}. \tag{6}$$

When $\frac{1}{2} < \nu < 1$, (6) is equivalent to $\lambda > 1/(\nu-\frac{1}{2})$. In terms of the moments of X, the latter assumption is equivalent to the finiteness of $\mathrm{E}(|X|^r)$ for some $r > 1/(\nu-\frac{1}{2})$. Furthermore, when $1 \leq \nu < \infty$, (6) is equivalent to $\lambda > 2$ which, in terms of the moments of X, is equivalent to $\mathrm{E}(X^2) < \infty$. Thus the assumptions of Theorem 2 are optimal in the Pareto case and cannot be improved upon in the general case (i.e. for arbitrary cdf F).

5. Estimation of $\sigma^2_{F,\nu}$ from a sample

In this section we discuss methods for estimating the asymptotic variance $\sigma^2_{F,\nu}$. If the cdf F is given in a parametric form, then $\sigma^2_{F,\nu}$ is a function of unknown parameters of F. Replacing the parameters by their maximum likelihood or other estimators, we obtain the corresponding estimators of $\sigma^2_{F,\nu}$. If F is not given in a parametric form, a non-parametric estimator of $\sigma^2_{F,\nu}$ can be constructed and used. For example, we construct a non-parametric estimator of $\sigma^2_{F,\nu}$ by replacing F everywhere in the definition of $\sigma^2_{F,\nu}$ by its empirical counterpart F_n. This leads us to the estimator

$$s^2_{n,\nu} = \frac{1}{\bar{X}^2}\big(s_n(\nu,\nu) + 2(I_{n,\nu}-1)s_n(1,\nu) + (I_{n,\nu}-1)^2 s_n(1,1)\big)$$

of $\sigma^2_{F,\nu}$ in which $s_n(\alpha, \beta)$ is a non-parametric estimator of $\sigma_F(\alpha, \beta)$, given by

$$s_n(\alpha, \beta) = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \phi_{ij}^{(n)}(\alpha, \beta)(X_{i+1:n} - X_{i:n})(X_{j+1:n} - X_{j:n}),$$

where $$\phi_{ij}^{(n)}(\alpha, \beta) = \alpha\beta\Big(\Big(\frac{i}{n} \wedge \frac{j}{n}\Big) - \frac{i}{n}\frac{j}{n}\Big)\Big(1 - \frac{i}{n}\Big)^{\alpha-1}\Big(1 - \frac{j}{n}\Big)^{\beta-1}.$$

To use the estimator $s^2_{n,\nu}$ for constructing confidence intervals for $I_{F,\nu}$, we need to demonstrate that $s^2_{n,\nu}$ is a consistent estimator of $\sigma^2_{F,\nu}$. Hence, the following theorem.

Theorem 3. *If the assumptions of Theorem 2 are satisfied, then $s^2_{n,\nu}$ is a strongly consistent estimator of $\sigma^2_{F,\nu}$.*

From Theorems 2 and 3, we obtain the approximate $100(1 - \alpha)\%$ confidence interval $I_{n,\nu} \pm z_{\alpha/2}\, s_{n,\nu}/\sqrt{n}$ for the S-Gini index $I_{F,\nu}$. The confidence interval provides a consistency check for confidence intervals obtained by the bootstrap methodology (cf. Xu, 2000).

The proof of Theorem 3 is technically complex and not particularly interesting in the context of this paper, so we do not present it here.

6. Recent developments

Since the submission of this paper for publication, we have learned about a closely related and very interesting paper by Barrett & Donald (2000). Using the empirical and quantile processes approach, Barrett and Donald develop a general large sample asymptotic theory for various indices of inequality and poverty, including the S-Gini index. Their investigations are mainly based on the fact that indices of inequality and poverty are functionals of Lorenz curves, and the Lorenz curves are (cf. Gastwirth, 1971) integrals of the corresponding quantile functions. Naturally then, using appropriate limit theorems for quantile processes, Barrett & Donald (2000) derive desired asymptotic results for Lorenz processes and, in turn, for the indices of inequality and poverty. In particular, their approach enables improvement on widely used methods based on considering only a finite number of Lorenz ordinates. For more details on the empirical processes approach and references, see e.g. the monograph by Csörgő, Csörgő & Horváth (1986). For further developments and more recent references concerning Lorenz curves see, e.g., Csörgő, Gastwirth & Zitikis (1998).

In this paper we treat the integral and the quantile function F^{-1} in

$$I_{F,\nu} = 1 - \frac{\nu}{\mu}\int_0^1 F^{-1}(t)(1 - t)^{\nu-1}\,dt$$

together, while Barrett & Donald (2000) treat them separately. This is the main difference between the two approaches. The advantage of our approach is that the generally 'bad' behaviour of the empirical quantile function F_n^{-1} near the two ends of the interval $(0, 1)$ is lessened by the operation of integration. Therefore, using this approach together with the theory of L-statistics, we avoid the assumptions made by Barrett & Donald (2000) (i.e. twice differentiability and compact support of F) and obtain consistency and asymptotic normality of the S-Gini index under only moment assumptions.

References

BARRETT, G.F. & DONALD, S.G. (2000). Statistical inference with generalized Gini indices of inequality and poverty. Discussion paper 2002/01. Sydney: School of Economics, University of New South Wales.

BLACKORBY, C., BOSSERT, W. & DONALDSON, D. (1999). Income inequality measurement: the normative approach (with a comment by J.A. Weymark). In *Handbook on Income Inequality Measurement*, ed. J. Silber, *Recent Economic Thought* **71**, pp. 133–161. Boston: Kluwer Academic Publishers.

CHAKRAVARTY, S.R. (1988). Extended Gini indices of inequality. *Internat. Econom. Rev.* **29**, 147–156.

CHERNOFF, H., GASTWIRTH, J.L. & JOHNS, M.V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* **38**, 52–72.

COWELL, F. A. (1999). Estimation of inequality indices (with a comment by A. Lemmi). In *Handbook on Income Inequality Measurement*, ed. J. Silber, *Recent Economic Thought* **71**, pp. 269–289. Boston: Kluwer Academic Publishers.

CSÖRGŐ, M., CSÖRGŐ, S. & HORVÁTH, L. (1986). *An Asymptotic Theory for Empirical Reliability and Concentration Processes.* Berlin: Springer.

CSÖRGŐ, M., GASTWIRTH, J.L. & ZITIKIS, R. (1998). Asymptotic confidence bands for the Lorenz and Bonferroni curves based on the empirical Lorenz curve. *J. Statist. Plann. Inference* **74**, 65–91.

DAVID, H.A. (1968). Gini's mean difference rediscovered. *Biometrika* **55**, 573–575.

DAVID, H.A. (1970). *Order Statistics.* New York: Wiley.

GASTWIRTH, J.L. (1971). A general definition of the Lorenz curve. *Econometrica* **39**, 1037–1039.

GIORGI, G.M. (1999). Income inequality measurement: the statistical approach (with a comment by M-P. Victoria-Feser). In *Handbook on Income Inequality Measurement*, ed. J. Silber, *Recent Economic Thought* **71**, pp. 245–267. Boston: Kluwer Academic Publishers.

GIORGI, G.M. & PALLINI, A. (1990). Inequality indices: theoretical and empirical aspects of their asymptotic behaviour. *Statist. Papers/Statistische Hefte* **31**, 65–76.

MEHRAN, F. (1976). Linear measures of income inequality. *Econometrica* **44**, 805–809.

NYGÅRD, F. & SANDSTRÖM, A. (1989). Income inequality measures based on sample surveys. *J. Econometrics* **42**, 81–95.

RYU, H.K. & SLOTTJE, D.J. (1999). Parametric approximations of the Lorenz curve (with a comment by J. Esteban). In *Handbook on Income Inequality Measurement*, ed. J. Silber, *Recent Economic Thought* **71**, pp. 291–314. Boston: Kluwer Academic Publishers.

SEN, A. (1997). *On Economic Inequality.* (Expanded edition with a substantial annexe by J.E. Foster and A. Sen.) Oxford: Clarendon Press.

SHORACK, G.R. (2000). *Probability for Statisticians.* New York: Springer.

SHORACK, G.R. & WELLNER, J.A. (1986). *Empirical Processes with Applications to Statistics.* New York: John Wiley & Sons.

SILBER, J. (Ed.) (1999). *Handbook on Income Inequality Measurement. Recent Economic Thought* **71**. Boston: Kluwer Academic Publishers.

WEYMARK, J.A. (1980/1981). Generalized Gini inequality indices. *Math. Social Sci.* **1**, 409–430.

XU, K. (2000). Inference for generalized Gini indices using the iterated-bootstrap method. *J. Bus. Econom. Statist.* **18**, 223–227.

YITZHAKI, S. (1983). On an extension of the Gini inequality index. *Internat. Econom. Rev.* **24**, 617–628.